Estimating the Markov-Switching Almost Ideal Demand Systems : A Bayesian approach (A New Perspective to Statistical Models and Its Related Topics)

KABE, Satoshi; KANAZAWA, Yuichiro

数理解析研究所講究録
第 1804 巻 2012 年 127-143

Estimating the Markov-Switching Almost Ideal Demand Systems: A Bayesian approach

Satoshi KABE [1]
Department of Social Systems and Management,
Graduate School of Systems and Information,
University of Tsukuba.

Yuichiro KANAZAWA [2]
Graduate School of Systems and Information,
University of Tsukuba.

1 Introduction

In empirical demand analysis, the almost ideal demand system (AIDS) model proposed by Deaton and Muellbauer (1980) has been widely used. This model assumes that market demand can be thought of as if it were the outcome of decisions by a rational representative consumer. In the previous studies on the structural shifts in demand, Moosa and Baxter (2002) developed the time-varying coefficient AIDS model. They introduced the stochastic trend and seasonality terms into the linear approximated AIDS (LA-AIDS) model so that it can be applied to unstable demand structure in the alcoholic beverage market in the U.K. Ishida et al. (2006, 2010) employed the gradual switching AIDS model. They set the transition functions into the AIDS model to capture the gradual shifts following Bovine Spongiform Encephalopathy (henceforth BSE) and bird flu outbreaks in the Japanese meat market. The latter model assumes that researchers know the structural change points in demand in advance.

Obviously modeling abrupt changes in demand caused by the unique exogenous events and estimating these change points are the next step. Allais and Nichèle (2007) seem to be the first to propose a Markov-switching almost ideal demand system (MS-AIDS) model extending the idea of Hamilton (1989) [3]. This model enables us to determine when the regime shifts occurred and to estimate parameters characterized across the different regimes. Moreover, degree of belongingness to each of the regimes and transitions between regimes are quantified by the probabilities. They analyzed the French meat and fish demands over the period 1991 - 2002 and detected the abrupt changes due to the two BSE crises in France. Kabe and Kanazawa (2012) also assessed the structural change points in the Japanese meat market during 1998 - 2006 via MS-AIDS model. They found the structural change point coinciding with the timing of first reported case of BSE, but not of bird flu. In both of these instances, MS-AIDS model is found to be quite effective in detecting abrupt changes in demand using monthly aggregate data.

In Allais and Nichèle (2007), they estimate the parameters including transition probabilities via maximum likelihood (ML) estimation. However, when the variance-covariance

---

[1]E-mail: k0420214@sk.tsukuba.ac.jp.

[2]Professor of Statistics, Email: kanazawa@sk.tsukuba.ac.jp.

[3]Hamilton (1989) proposed the Markov-switching model to date the timing of recessions and booms with real gross national product (GNP) data in U.S. He found that the regime shift from positive to negative growth rate has a recurrent feature of the U.S. business cycle.

matrices differ between regimes, a singularity problem arises when the determinant of variance-covariance matrix is close to zero, sending log-likelihood of MS-AIDS model to infinity, and making numerical optimization methods (e.g., Newton-Raphson method) break down. This problem is well-known in the literature on estimation of mixture of normal distributions. They also estimate the transition probabilities via ML estimation without any constraints, although the transition probabilities have to lie between zero and one inclusive.

To avoid the singularity problem, subjective judgment is required in deciding what constitutes a suitable region for plausible value of the variance-covariance matrices, so that Hamilton (1990, 1991) suggests the Bayesian estimation as a simple solution of the singularity problem. Hamilton (1991, p.37) stated that "the [Bayesian] approach is intuitively appealing and trivial to implement. Monte Carlo analysis suggests that this approach can consistently improve the MSE's for a wide variety of underlying models."

Bayesian estimation enables us to incorporate the prior information on the variance-covariance matrices to the conjugate prior distributions. Moreover, Bayesian estimation can provide us with the posterior distributions of transition probabilities in the unit intervals, avoiding problems associated with unconstrained ML estimation. Finally, Bayesian estimation is able to replace the messy calculations entailed in the score functions of log-likelihood for MS-AIDS model with computationally simple Gibbs sampler. To the best of our knowledge, no Bayesian estimation method is proposed to solve theses problems associated with the MS-AIDS model. In this paper, we propose a Bayesian method to estimate parameters in MS-AIDS model along with the transition probabilities. To illustrate its applicability, we take the proposed method to the Japanese meat market data and examine the regime shifts caused by the food safety concerns such as BSE and bird flu.

The rest of this paper is organized as follows. Section 2 briefly describes the Markov-switching AIDS model and introduces the necessary notations. Then section 3 proposes the Bayesian estimation, and section 4 presents the empirical study on the Japanese meat market via the proposed Bayesian estimation method. Finally, section 5 discusses the merits of the proposed Bayesian method relative to the ML estimation we employed in Kabe and Kanazawa (2012), and then we point to future directions of the research.

2 Markov-Switching AIDS model

Suppose that s_t is an unobserved random variable that takes an integer value in $1, 2, \ldots, K$ to express "regime" or "state" at time t, then budget share of i-th product at time t, $\bar{w}_{it}$ which is defined as $p_{it}q_{it}/m_{0t}$ with price p_{it}, quantity q_{it} and expenditure (or budget) m_{0t} $(= \sum_i p_{it}q_{it})$ takes the following form:

$$\bar{w}_{it} = \alpha_{i,s_t} + \sum_{j=1}^{N} \gamma_{ij,s_t} \log p_{jt} + \beta_{i,s_t} \log\left(\frac{m_{0t}}{P_t}\right) \tag{2.1}$$

where P_t is a price index which is defined by

$$\log P_t = \alpha_{0,s_t} + \sum_{k=1}^{N} \alpha_{k,s_t} \log p_{kt} + \frac{1}{2}\sum_{k=1}^{N}\sum_{j=1}^{N} \gamma_{kj,s_t} \log p_{kt} \log p_{jt} \tag{2.2}$$

and α_{0,s_t}, α_{i,s_t}, γ_{ij,s_t} and β_{i,s_t} $(i,j=1,2,\dots,N)$ are regime-dependent parameters.

The parameters in (2.1) and (2.2) have the theoretical constraints [4] as follows

$$\text{[Adding up]} \qquad \sum_{i=1}^{N}\alpha_{i,s_t}=1,\ \sum_{i=1}^{N}\gamma_{ij,s_t}=0,\ \sum_{i=1}^{N}\beta_{i,s_t}=0, \tag{2.3a}$$

$$\text{[Homogeneity]} \qquad \sum_{j=1}^{N}\gamma_{ij,s_t}=0, \tag{2.3b}$$

$$\text{[Symmetry]} \qquad \gamma_{ij,s_t}=\gamma_{ji,s_t}. \tag{2.3c}$$

Following the previous studies (Rickertsen, 1996; Allais and Nichèle, 2007; Ishida et al., 2010), we include a trend effect, seasonal effect and habit effect into the intercept term α_{i,s_t} as

$$\alpha_{i,s_t}=\bar{\alpha}_{i,s_t}+\nu_{i,s_t}t+\delta_{1,i}d_{1,t}+\delta_{2,i}d_{2,t}+\sum_{j=1}^{N}\phi_{ij}\bar{w}_{j,t-1} \tag{2.4}$$

where $d_{1,t}$ and $d_{2,t}$ are dummy variables

$$d_{1,t}=\begin{cases}1 & \text{if } t \text{ is August}\\ 0 & \text{otherwise}\end{cases} \qquad d_{2,t}=\begin{cases}1 & \text{if } t \text{ is December}\\ 0 & \text{otherwise.}\end{cases}$$

As for seasonal effect, we set the dummy variables to adjust the seasonality in budget shares. The budget shares for meat and fish are considered to shift due to the seasonal habits (e.g., summer camp, gift-giving tradition, year-end party and so forth) in August and December. Furthermore, we include a habit effect which is defined as a linear function of one-lagged budget shares (Rickertsen, 1996; Allais and Nichèle, 2007). In order to satisfy the adding up condition, we impose the restriction $\sum_{i=1}^{N}\bar{\alpha}_{i,s_t}=1$, $\sum_{i=1}^{N}\nu_{i,s_t}=0$, $\sum_{i=1}^{N}\delta_{1,i}=\sum_{i=1}^{N}\delta_{2,i}=0$ and $\sum_{i=1}^{N}\phi_{ij}=0$. We also impose the restriction $\sum_{j=1}^{N}\phi_{ij}=0$ to avoid the identification problem.

Using the theoretical constraints in (2.3a), (2.3b) and (2.3c), the MS-AIDS model (2.1) can be rewritten as

$$\bar{w}_{it}=\alpha_{i,s_t}+\sum_{j=1}^{N-1}\gamma_{ij,s_t}\log\left(\frac{p_{jt}}{p_{Nt}}\right)+\beta_{i,s_t}\log\left(\frac{m_{0t}}{P_t}\right) \tag{2.5}$$

where $i=1,2,\dots,N-1$. Imposing the restriction $\sum_{j=1}^{N}\phi_{ij}=0$, intercept term α_{i,s_t} in (2.5) is expressed as

$$\alpha_{i,s_t}=\bar{\alpha}_{i,s_t}+\nu_{i,s_t}t+\delta_{1,i}d_{1,t}+\delta_{2,i}d_{2,t}+\sum_{j=1}^{N-1}\phi_{ij}(\bar{w}_{j,t-1}-\bar{w}_{N,t-1}).$$

[4] "Adding up" guarantees that the total expenditure is equal to the sum of expenditures on the category of products under consideration. "Homogeneity" guarantees that if prices of products increase to $\tau p_{1t},\dots,\tau p_{Nt}$ for a scalar $\tau>0$, representative consumer has to increase his expenditure from m_{0t} to τm_{0t} to keep his utility level. "Symmetry" guarantees that the substitution effect in the Slutsky equation is symmetric.

The MS-AIDS model employs the Markov switching mechanism which is developed by Hamilton (1989). The Markov switching mechanism can express switching of regimes by using the unobserved random variables that follow the Markov process. To apply the Markov switching mechanism, we assume that transitions between regimes are governed by a K-state Markov chain with transition probabilities:

$$\Pr(s_t = j | s_{t-1} = i) = \pi_{ij}, \quad i, j = 1, 2, \ldots, K \tag{2.6}$$

and the transition matrix is defined as

$$\boldsymbol{\Pi} = \begin{bmatrix} \pi_{11} & \pi_{21} & \cdots & \pi_{K1} \\ \pi_{12} & \pi_{22} & \cdots & \pi_{K2} \\ \vdots & \vdots & \ddots & \vdots \\ \pi_{1K} & \pi_{2K} & \cdots & \pi_{KK} \end{bmatrix} \tag{2.7}$$

where $\pi_{i1} + \pi_{i2} + \cdots + \pi_{iK} = 1$, $i = 1, 2, \ldots, K$.

3 Bayesian Estimation

Let $\boldsymbol{w}_t$ be a $(N-1) \times 1$ vector of budget shares at time t, $\bar{w}_{it}$ $(i = 1, 2, \ldots, N-1)$ and we define the matrix of explanatory variables for regime-dependent parameters $\bar{\alpha}_{i,s_t}$, γ_{i1,s_t}, γ_{i2,s_t}, $\ldots$, γ_{iN-1,s_t}, β_{i,s_t}, ν_{i,s_t} as $\boldsymbol{X}_t^{(1)}$ and for regime-independent parameters $\delta_{1,i}$, $\delta_{2,i}$, ϕ_{i1}, ϕ_{i2}, $\ldots$, $\phi_{i,N-1}$ as $\boldsymbol{X}_t^{(0)}$.

Given the value of price index (2.2), the MS-AIDS model (2.5) can be first rewritten by separating the parts that depend on regimes and by including the error term ε_{it} as

$$\begin{aligned} \bar{w}_{it} = \bar{\alpha}_{i,s_t} &+ \sum_{j=1}^{N-1} \gamma_{ij,s_t} \log\left(\frac{p_{jt}}{p_{Nt}}\right) + \beta_{i,s_t} \log\left(\frac{m_{0t}}{P_t}\right) + \nu_{i,s_t} t \\ &+ \delta_{1,i} d_{1,t} + \delta_{2,i} d_{2,t} + \sum_{j=1}^{N-1} \phi_{ij} (\bar{w}_{j,t-1} - \bar{w}_{N,t-1}) + \varepsilon_{it} \end{aligned} \tag{3.1}$$

and thus can further be rewritten as the matrix form:

$$\boldsymbol{w}_t = \boldsymbol{X}_t^{(1)} \boldsymbol{\theta}_{s_t} + \boldsymbol{X}_t^{(0)} \boldsymbol{\theta}_0 + \boldsymbol{\varepsilon}_t \tag{3.2}$$

where $\boldsymbol{\varepsilon}_t \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_{s_t})$ and $\boldsymbol{\Sigma}_{s_t}$ is also regime-dependent parameter such that $\boldsymbol{\Sigma}_{s_t} = \boldsymbol{\Sigma}_j$ if time t belongs to regime j. The size of the matrices $\boldsymbol{X}_t^{(1)}$ and $\boldsymbol{X}_t^{(0)}$ are $(N-1) \times [\,3(N-1) + N(N-1)/2\,]$ and $(N-1) \times (N-1)(N+1)$.

Example

Let us consider the case that the number of products N is four. Then 3×15 matrix $\boldsymbol{X}_t^{(1)}$ is defined as

$$\boldsymbol{X}_t^{(1)} \equiv \begin{bmatrix} \boldsymbol{I}_3 & \boldsymbol{P}_t & \boldsymbol{M}_t & \boldsymbol{T}_t \end{bmatrix}$$

where $\boldsymbol{I}_3$ is a 3×3 identity matrix,

$$\boldsymbol{M}_t \equiv \begin{bmatrix} \log(\frac{m_{0t}}{P_t}) & & 0 \\ & \log(\frac{m_{0t}}{P_t}) & \\ 0 & & \log(\frac{m_{0t}}{P_t}) \end{bmatrix}, \quad \boldsymbol{T}_t \equiv \begin{bmatrix} t & & 0 \\ & t & \\ 0 & & t \end{bmatrix},$$

and

$$\boldsymbol{P}_t \equiv \begin{bmatrix} \log(\frac{p_{1t}}{p_{4t}}) & \log(\frac{p_{2t}}{p_{4t}}) & \log(\frac{p_{3t}}{p_{4t}}) & 0 & 0 & 0 \\ 0 & \log(\frac{p_{1t}}{p_{4t}}) & 0 & \log(\frac{p_{2t}}{p_{4t}}) & \log(\frac{p_{3t}}{p_{4t}}) & 0 \\ 0 & 0 & \log(\frac{p_{1t}}{p_{4t}}) & 0 & \log(\frac{p_{2t}}{p_{4t}}) & \log(\frac{p_{3t}}{p_{4t}}) \end{bmatrix}.$$

The 15 element parameter vector $\boldsymbol{\theta}_{s_t}$ is defined as

$$\boldsymbol{\theta}_{s_t} \equiv \begin{bmatrix} \bar{\boldsymbol{\alpha}}_{s_t} \\ \boldsymbol{\gamma}_{s_t} \\ \boldsymbol{\beta}_{s_t} \\ \boldsymbol{\nu}_{s_t} \end{bmatrix}$$

where $\bar{\boldsymbol{\alpha}}_{s_t} \equiv [\ \bar{\alpha}_{1,s_t}\ \bar{\alpha}_{2,s_t}\ \bar{\alpha}_{3,s_t}\ ]'$, $\boldsymbol{\gamma}_{s_t} \equiv [\ \gamma_{11,s_t}\ \gamma_{12,s_t}\ \gamma_{13,s_t}\ \gamma_{22,s_t}\ \gamma_{23,s_t}\ \gamma_{33,s_t}\ ]'$, $\boldsymbol{\beta}_{s_t} \equiv [\ \beta_{1,s_t}\ \beta_{2,s_t}\ \beta_{3,s_t}\ ]'$ and $\boldsymbol{\nu}_{s_t} \equiv [\ \nu_{1,s_t}\ \nu_{2,s_t}\ \nu_{3,s_t}\ ]'$.

The 3×15 matrix $\boldsymbol{X}_t^{(0)}$ is defined as

$$\boldsymbol{X}_t^{(0)} \equiv \begin{bmatrix} \boldsymbol{D}_{1t} & \boldsymbol{D}_{2t} & \boldsymbol{W}_{1t} & \boldsymbol{W}_{2t} & \boldsymbol{W}_{3t} \end{bmatrix}$$

where

$$\boldsymbol{D}_{1t} \equiv \begin{bmatrix} d_{1,t} & & 0 \\ & d_{1,t} & \\ 0 & & d_{1,t} \end{bmatrix}, \quad \boldsymbol{D}_{2t} \equiv \begin{bmatrix} d_{2,t} & & 0 \\ & d_{2,t} & \\ 0 & & d_{2,t} \end{bmatrix},$$

and

$$\boldsymbol{W}_{jt} \equiv \begin{bmatrix} \bar{w}_{j,t-1} - \bar{w}_{4,t-1} & & 0 \\ & \bar{w}_{j,t-1} - \bar{w}_{4,t-1} & \\ 0 & & \bar{w}_{j,t-1} - \bar{w}_{4,t-1} \end{bmatrix}.$$

The 15 element parameter vector $\boldsymbol{\theta}_0$ is defined as

$$\boldsymbol{\theta}_0 \equiv \begin{bmatrix} \boldsymbol{\delta}_1 \\ \boldsymbol{\delta}_2 \\ \boldsymbol{\phi}_1 \\ \boldsymbol{\phi}_2 \\ \boldsymbol{\phi}_3 \end{bmatrix}$$

where $\boldsymbol{\delta}_1 \equiv [\ \delta_{11}\ \delta_{12}\ \delta_{13}\ ]'$, $\boldsymbol{\delta}_2 \equiv [\ \delta_{21}\ \delta_{22}\ \delta_{23}\ ]'$ and $\boldsymbol{\phi}_j \equiv [\ \phi_{1j}\ \phi_{2j}\ \phi_{3j}]'$.

Since the MS-AIDS model is parameterized nonlinear due to the price index (2.2), the estimate cannot be written as closed form. Instead, we assume that value of price index is already known and draw samples of parameters via Gibbs sampler. Afterward, these samples are used to update the value of price index and then samples of parameters are generated with the new price index. We repeat this process until convergence.

To obtain the likelihood function of MS-AIDS model, we denote the set of variables obtained from $t = 1$ through time t as

$$\mathcal{Y}_t \equiv \{\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_t\}, \quad \mathcal{S}_t \equiv \{s_1, s_2, \ldots, s_t\},$$
$$\mathcal{X}_t \equiv \{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_t\}$$

where $\boldsymbol{x}_t$ is a $1 \times (2N+3)$ vector of explanatory variables at time t in (3.1): $\boldsymbol{x}_t \equiv [\ 1\ \log(p_{1t}/p_{Nt})\ \cdots\ \log(p_{N-1t}/p_{Nt})\ \log(m_{0t}/P_t)\ t\ d_{1,t}\ d_{2,t}\ \bar{w}_{1,t-1} - \bar{w}_{N,t-1}\ \cdots\ \bar{w}_{N-1,t-1} - \bar{w}_{N,t-1}]$.

Then likelihood function $\mathcal{L}(\cdot|\cdot)$ is defined as

$$\mathcal{L}(\boldsymbol{\theta}, \boldsymbol{\pi}|\mathcal{Y}_T, \mathcal{S}_T, \mathcal{X}_T) = \mathcal{L}(\boldsymbol{\pi}|\mathcal{S}_T)\mathcal{L}(\boldsymbol{\theta}|\mathcal{Y}_T, \mathcal{S}_T, \mathcal{X}_T) \tag{3.3}$$

where $\boldsymbol{\theta} \equiv \{\boldsymbol{\theta}_0, \boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \ldots, \boldsymbol{\theta}_K, \Sigma_1, \Sigma_2, \ldots, \Sigma_K\}$ and $\boldsymbol{\pi} \equiv \{\pi_{ij} : i, j = 1, 2, \ldots, K\}$. Given a prior distribution $p(\boldsymbol{\theta}, \boldsymbol{\pi}) = p(\boldsymbol{\theta})p(\boldsymbol{\pi})$ [5], we obtain the posterior distributions with respect to $\boldsymbol{\theta}$ and to $\boldsymbol{\pi}$ as

$$\begin{aligned} p(\boldsymbol{\theta}, \boldsymbol{\pi}|\mathcal{Y}_T, \mathcal{S}_T, \mathcal{X}_T) &\propto \mathcal{L}(\boldsymbol{\theta}, \boldsymbol{\pi}|\mathcal{Y}_T, \mathcal{S}_T, \mathcal{X}_T)p(\boldsymbol{\theta}, \boldsymbol{\pi}) \\ &= \mathcal{L}(\boldsymbol{\pi}|\mathcal{S}_T)p(\boldsymbol{\pi}) \times \mathcal{L}(\boldsymbol{\theta}|\mathcal{Y}_T, \mathcal{S}_T, \mathcal{X}_T)p(\boldsymbol{\theta}) \\ &\propto p(\boldsymbol{\pi}|\mathcal{S}_T) \times p(\boldsymbol{\theta}|\mathcal{Y}_T, \mathcal{S}_T, \mathcal{X}_T). \end{aligned} \tag{3.4}$$

Now we compute each of the terms on the right hand side of (3.4). To be able to do this, we first need to generate discrete latent variables $s_1, s_2, \ldots, s_T$ to represent regimes.

3.1 Sampling of latent variables $s_1, s_2, \ldots, s_T$

Since $\mathcal{S}_T \equiv \{s_1, s_2, \ldots, s_T\}$ is a sequence of unobservable finite discrete random variables, we need to generate samples $s_1, s_2, \ldots, s_T$ to compute the posterior distributions in (3.4). To generate samples of latent variables $s_1, s_2, \ldots, s_T$, we apply the multi-move sampler (e.g., Carter and Kohn, 1994; Chib, 1996): Given the data obtained through time t, $\boldsymbol{\Omega}_t \equiv \{\mathcal{Y}_t, \mathcal{X}_t\}$ and set of parameters $\boldsymbol{\Theta} \equiv \{\boldsymbol{\theta}, \boldsymbol{\pi}\}$, we consider a joint distribution $f(\mathcal{S}_T|\boldsymbol{\Omega}_T, \boldsymbol{\Theta})$ as

$$\begin{aligned} f(\mathcal{S}_T|\boldsymbol{\Omega}_T, \boldsymbol{\Theta}) &= f(s_1, s_2, \ldots, s_T|\boldsymbol{\Omega}_T, \boldsymbol{\Theta}) \\ &= \Pr(s_T|\boldsymbol{\Omega}_T, \boldsymbol{\Theta}) \Pr(s_{T-1}|s_T, \boldsymbol{\Omega}_{T-1}, \boldsymbol{\Theta}) \cdots \Pr(s_1|s_2, \boldsymbol{\Omega}_1, \boldsymbol{\Theta}) \\ &= \Pr(s_T|\boldsymbol{\Omega}_T, \boldsymbol{\Theta}) \prod_{t=1}^{T-1} \Pr(s_t|s_{t+1}, \boldsymbol{\Omega}_t, \boldsymbol{\Theta}) \end{aligned} \tag{3.5}$$

if $\mathcal{S}_T \equiv \{s_1, s_2, \ldots, s_T\}$ is assumed to follow Markov process. This usage of Markovian in reverse order is justified by virtue of Bayes theorem as we see below

$$\begin{aligned} \Pr(s_t|s_{t+1}, \boldsymbol{\Omega}_t, \boldsymbol{\Theta}) &= \frac{\Pr(s_{t+1}|s_t) \Pr(s_t|\boldsymbol{\Omega}_t, \boldsymbol{\Theta})}{\Pr(s_{t+1}|\boldsymbol{\Omega}_t, \boldsymbol{\Theta})} \\ &= \frac{\Pr(s_{t+1}|s_t) \Pr(s_t|\boldsymbol{\Omega}_t, \boldsymbol{\Theta})}{\sum_{s_t=1}^{K} \Pr(s_{t+1}|s_t) \Pr(s_t|\boldsymbol{\Omega}_t, \boldsymbol{\Theta})} \end{aligned} \tag{3.6}$$

where $\Pr(s_{t+1}|s_t)$ is a transition probability. Notice that $\Pr(s_t|s_{t+1}, \boldsymbol{\Omega}_t, \boldsymbol{\Theta})$ in (3.5) can be computed from (3.6). The quantity $\Pr(s_t|\boldsymbol{\Omega}_t, \boldsymbol{\Theta})$ can be derived by using the Hamilton filter (Hamilton, 1989).

[5] That is, the prior of $\boldsymbol{\theta}$ and the prior of $\boldsymbol{\pi}$ are independent.

3.2 Sampling of transition probabilities π_{ij}

Given the samples of latent variables $s_1, s_2, \ldots, s_T$, the likelihood function $\mathcal{L}(\boldsymbol{\pi}|\mathcal{S}_T)$ appears on the right hand side of (3.4) is defined as

$$\mathcal{L}(\boldsymbol{\pi}|\mathcal{S}_T) = \prod_{i=1}^{K}\prod_{j=1}^{K} \pi_{ij}^{n_{ij}}$$

where n_{ij} is the total number of transitions from i to j from $t = 1$ to $t = T$.

Suppose that the i-th column vector of transition matrix (2.7) is denoted by $\boldsymbol{\pi}_i = [\pi_{i1}\ \pi_{i2}\ \cdots\ \pi_{iK}]'$ and let the prior distribution of $\boldsymbol{\pi}_i$, independently of $\boldsymbol{\pi}_j$ $(j \neq i)$ be a K-dimensional Dirichlet distribution [6] :

$$\boldsymbol{\pi}_i \sim Dir(u_{i1}, u_{i2}, \ldots, u_{iK}),$$

then posterior distribution of $\boldsymbol{\pi}_i$ is given as

$$\begin{aligned} p(\boldsymbol{\pi}_i|\mathcal{S}_T) &\propto \mathcal{L}(\boldsymbol{\pi}_i|\mathcal{S}_T)p(\boldsymbol{\pi}_i) \\ &\propto \prod_{j=1}^{K} \pi_{ij}^{n_{ij}} \times (\pi_{i1}^{u_{i1}-1} \cdots \pi_{iK}^{u_{iK}-1}) \\ &= \pi_{i1}^{n_{i1}+u_{i1}-1}\pi_{i2}^{n_{i2}+u_{i2}-1} \cdots \pi_{iK}^{n_{iK}+u_{iK}-1}. \end{aligned} \tag{3.7}$$

Therefore we have

$$\boldsymbol{\pi}_i|\mathcal{S}_T \sim Dir(n_{i1}+u_{i1}, n_{i2}+u_{i2}, \ldots, n_{iK}+u_{iK}), \quad i = 1, 2, \ldots, K.$$

This corresponds to the first term on the right hand side of (3.4).

3.3 Sampling of parameters $\boldsymbol{\theta}_0$, $\boldsymbol{\theta}_j$, $j = 1, 2, \ldots, K$

To evaluate the posterior distribution $p(\boldsymbol{\theta}|\mathcal{Y}_T, \mathcal{S}_T, \mathcal{X}_T)$ in (3.4) via Gibbs sampler, we need conditional distributions of $\{\boldsymbol{\theta}_j\}_{j=0}^K$ given $\{\Sigma_j\}_{j=1}^K$ and of $\{\Sigma_j\}_{j=1}^K$ given $\{\boldsymbol{\theta}_j\}_{j=0}^K$. To make the formulation clear, we assume that prior distribution $p(\boldsymbol{\theta})$ [7] can be written as

$$\begin{aligned} p(\boldsymbol{\theta}) &\equiv p(\{\boldsymbol{\theta}_j\}_{j=0}^K, \{\Sigma_j\}_{j=1}^K) \\ &= p(\{\boldsymbol{\theta}_j\}_{j=0}^K)p(\{\Sigma_j\}_{j=1}^K). \end{aligned}$$

Since posterior distribution $p(\boldsymbol{\theta}|\mathcal{Y}_T, \mathcal{S}_T, \mathcal{X}_T)$ in (3.4) can be rewritten as

$$p(\boldsymbol{\theta}|\mathcal{Y}_T, \mathcal{S}_T, \mathcal{X}_T) \equiv p(\{\boldsymbol{\theta}_j\}_{j=0}^K, \{\Sigma_j\}_{j=1}^K|\mathcal{Y}_T, \mathcal{S}_T, \mathcal{X}_T),$$

[6] The Dirichlet distribution for $\boldsymbol{\pi}_i$ is defined as

$$p(\boldsymbol{\pi}_i|u_{i1}, u_{i2}, \ldots, u_{iK}) = \frac{\Gamma(u_{i0})}{\Gamma(u_{i1})\cdots\Gamma(u_{iK})}\pi_{i1}^{u_{i1}-1}\cdots\pi_{iK}^{u_{iK}-1}$$

where $0 \leq \pi_{ij} \leq 1$, $\sum_{j=1}^K \pi_{ij} = 1$, $u_{ij} > 0$, and $u_{i0} = \sum_{j=1}^K u_{ij}$.

[7] In other words, we assume that the prior distribution's location parameters $\{\boldsymbol{\theta}_j\}_{j=0}^K$ and scale-like parameters $\{\Sigma_j\}_{j=1}^K$ can be freely moved and form a K-dimensional rectangular parameter space.

we have

$$p(\{\boldsymbol{\theta}_j\}_{j=0}^K, \{\boldsymbol{\Sigma}_j\}_{j=1}^K | \mathcal{Y}_T, \mathcal{S}_T, \mathcal{X}_T)$$
$$\propto \mathcal{L}(\{\boldsymbol{\theta}_j\}_{j=0}^K, \{\boldsymbol{\Sigma}_j\}_{j=1}^K | \mathcal{Y}_T, \mathcal{S}_T, \mathcal{X}_T) p(\{\boldsymbol{\theta}_j\}_{j=0}^K) p(\{\boldsymbol{\Sigma}_j\}_{j=1}^K). \tag{3.8}$$

Given $\{\boldsymbol{\Sigma}_j\}_{j=1}^K$, $\mathcal{Y}_T$, $\mathcal{S}_T$, and $\mathcal{X}_T$, conditional posterior distributions for $\{\boldsymbol{\theta}_j\}_{j=0}^K$ is expressed by dividing both sides of (3.8) by $p(\{\boldsymbol{\Sigma}_j\}_{j=1}^K)$

$$p(\{\boldsymbol{\theta}_j\}_{j=0}^K | \mathcal{Y}_T, \mathcal{S}_T, \mathcal{X}_T, \{\boldsymbol{\Sigma}_j\}_{j=1}^K)$$
$$\propto \mathcal{L}(\{\boldsymbol{\theta}_j\}_{j=0}^K, \{\boldsymbol{\Sigma}_j\}_{j=1}^K | \mathcal{Y}_T, \mathcal{S}_T, \mathcal{X}_T) p(\{\boldsymbol{\theta}_j\}_{j=0}^K). \tag{3.9}$$

Given the samples of latent variables $s_1, s_2, \ldots, s_T$ obtained in section 3.1 and variance-covariance matrices $\boldsymbol{\Sigma}_j$ $(j = 1, 2, \ldots, K)$ to be described in (3.13), we re-express the MS-AIDS model (3.2) as

$$\boldsymbol{w}_t = \boldsymbol{X}_t \boldsymbol{\theta}^* + \boldsymbol{\varepsilon}_t \tag{3.10}$$

where $\boldsymbol{\varepsilon}_t \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_{s_t})$. The matrix $\boldsymbol{X}_t$ in (3.10) is defined as

$$\boldsymbol{X}_t = \begin{bmatrix} \mathbf{1}\{s_t = 1\}\boldsymbol{X}_t^{(1)} & \mathbf{1}\{s_t = 2\}\boldsymbol{X}_t^{(1)} & \cdots & \mathbf{1}\{s_t = K\}\boldsymbol{X}_t^{(1)} & \boldsymbol{X}_t^{(0)} \end{bmatrix}$$

with indicator function $\mathbf{1}\{s_t = j\}$ taking scalar value 1 if $s_t = j$, or 0 otherwise [8] , and parameter vector $\boldsymbol{\theta}^*$ is defined as

$$\boldsymbol{\theta}^* \equiv \begin{bmatrix} \boldsymbol{\theta}_1 \\ \boldsymbol{\theta}_2 \\ \vdots \\ \boldsymbol{\theta}_K \\ \boldsymbol{\theta}_0 \end{bmatrix}.$$

To generate samples of $\boldsymbol{\theta}^*$, we derive the posterior distribution of $\boldsymbol{\theta}^*$ conditional on $\boldsymbol{\Sigma}_1$, $\boldsymbol{\Sigma}_2$, ..., $\boldsymbol{\Sigma}_K$ from (3.9). Applying the multivariate normal distribution $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{V})$ as a conjugate prior $p(\boldsymbol{\theta}^*)$, conditional posterior distribution of $\boldsymbol{\theta}^*$ in (3.9) is derived as

$$p(\boldsymbol{\theta}^* | \mathcal{Y}_T, \mathcal{S}_T, \mathcal{X}_T, \{\boldsymbol{\Sigma}_j\}_{j=1}^K)$$
$$\propto \prod_{t=1}^T \left[(2\pi)^{-\frac{N-1}{2}} |\boldsymbol{\Sigma}_{s_t}|^{-\frac{1}{2}} \exp\left\{ -\frac{1}{2}(\boldsymbol{w}_t - \boldsymbol{X}_t\boldsymbol{\theta}^*)' \boldsymbol{\Sigma}_{s_t}^{-1} (\boldsymbol{w}_t - \boldsymbol{X}_t\boldsymbol{\theta}^*) \right\} \right]$$
$$\times |\boldsymbol{V}|^{-\frac{1}{2}} \exp\left\{ -\frac{1}{2}(\boldsymbol{\theta}^* - \boldsymbol{\mu})' \boldsymbol{V}^{-1} (\boldsymbol{\theta}^* - \boldsymbol{\mu}) \right\}$$
$$\propto \exp\left\{ (\boldsymbol{\theta}^* - \boldsymbol{b})' \boldsymbol{B}^{-1} (\boldsymbol{\theta}^* - \boldsymbol{b}) \right\}$$

where

$$\boldsymbol{b} = \boldsymbol{B}\left(\sum_{t=1}^T \boldsymbol{X}_t' \boldsymbol{\Sigma}_{s_t}^{-1} \boldsymbol{w}_t + \boldsymbol{V}^{-1}\boldsymbol{\mu} \right), \quad \boldsymbol{B}^{-1} = \sum_{t=1}^T \boldsymbol{X}_t' \boldsymbol{\Sigma}_{s_t}^{-1} \boldsymbol{X}_t + \boldsymbol{V}^{-1}.$$

Then the conditional posterior distribution of $\boldsymbol{\theta}^*$ is

$$\boldsymbol{\theta}^* | \mathcal{Y}_T, \mathcal{S}_T, \mathcal{X}_T, \{\boldsymbol{\Sigma}_j\}_{j=1}^K \sim \mathcal{N}(\boldsymbol{b}, \boldsymbol{B}). \tag{3.11}$$

[8] That is, when $s_t = k$, the matrix $\boldsymbol{X}_t$ consists of $k-1$ of matrices of size $(N-1) \times [3(N-1) + N(N-1)/2]$ whose elements are all zero, $\boldsymbol{X}_t^{(1)}$, and $K-k$ of matrices of size $(N-1) \times [3(N-1) + N(N-1)/2]$ whose elements are all zero, and $\boldsymbol{X}_t^{(0)}$, all aligned from left to right.

3.4 Sampling of parameters Σ_j, $j = 1, 2, \ldots, K$

We assume that $\Sigma_1, \Sigma_2, \ldots, \Sigma_K$ are independent, then conditional posterior distribution for $\{\Sigma_j\}_{j=1}^K$ is expressed as

$$
\begin{aligned}
p(\{\Sigma_j\}_{j=1}^K | \mathcal{Y}_T, \mathcal{S}_T, \mathcal{X}_T, \{\boldsymbol{\theta}_j\}_{j=0}^K) &= \prod_{j=1}^K p(\Sigma_j | \mathcal{Y}_T, \mathcal{S}_T, \mathcal{X}_T, \{\boldsymbol{\theta}_j\}_{j=0}^K) \\
&\propto \prod_{j=1}^K \mathcal{L}(\Sigma_j, \{\boldsymbol{\theta}_j\}_{j=0}^K | \mathcal{Y}_T, \mathcal{S}_T, \mathcal{X}_T) p(\Sigma_j). \qquad (3.12)
\end{aligned}
$$

To generate samples of Σ_j, $j = 1, 2, \ldots, K$, we derive the conditional posterior distribution of Σ_j from (3.12). Applying the inverse Wishart distribution $\mathcal{IW}(\nu_j, \Lambda_j)$ as a conjugate prior $p(\Sigma_j)$, conditional posterior distribution of Σ_j is derived as

$$
\begin{aligned}
&p(\Sigma_j | \mathcal{Y}_T, \mathcal{S}_T, \mathcal{X}_T, \boldsymbol{\theta}^*) \\
&\propto \prod_{t \in \{t : s_t = j\}} \left[(2\pi)^{-\frac{N-1}{2}} |\Sigma_j|^{-\frac{1}{2}} \exp\left(-\frac{1}{2} \boldsymbol{\varepsilon}_t' \Sigma_j^{-1} \boldsymbol{\varepsilon}_t \right) \right] \\
&\quad \times |\Sigma_j|^{-\frac{\nu_j + (N-1) + 1}{2}} \exp\left(-\frac{1}{2} \mathrm{tr} \left\{ \Sigma_j^{-1} \Lambda_j \right\} \right) \\
&\propto |\Sigma_j|^{-\frac{\nu_j + (N-1) + 1 + n_j}{2}} \exp\left(-\frac{1}{2} \mathrm{tr} \left\{ \Sigma_j^{-1} \left(\sum_{t=1}^T \boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t' \mathbf{1}\{s_t = j\} + \Lambda_j \right) \right\} \right)
\end{aligned}
$$

where n_j is the total number of time t belonging to regime j. Then conditional posterior distribution of Σ_j is

$$
\Sigma_j | \mathcal{Y}_T, \mathcal{S}_T, \mathcal{X}_T, \boldsymbol{\theta}^* \sim \mathcal{IW}\left(\nu_j + n_j, \sum_{t=1}^T \boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t' \mathbf{1}\{s_t = j\} + \Lambda_j \right) \qquad (3.13)
$$

where $j = 1, 2, \ldots, K$.

From (3.11) and (3.13), we are able to construct Gibbs sampler algorithm by generating $\boldsymbol{\theta}^*$ and substituting these into (3.13) and then generating Σ_j with the generated $\boldsymbol{\theta}^*$ and substituting those back into (3.11).

4 Empirical study on Japanese meat market

The Ministry of Internal Affairs and Communications in Japan provides us with the household expenditure survey data (i.e., Family Income and Expenditure Survey). The household expenditure survey data includes the monthly time series data about average expenditure and price of meat and fish products along with others. In this study, we used the average expenditure and price data of beef, pork, chicken and fish over January 1998 to December 2006 (108 months). Figure 1 plots the budget shares of meat products from January 1998 through December 2006.

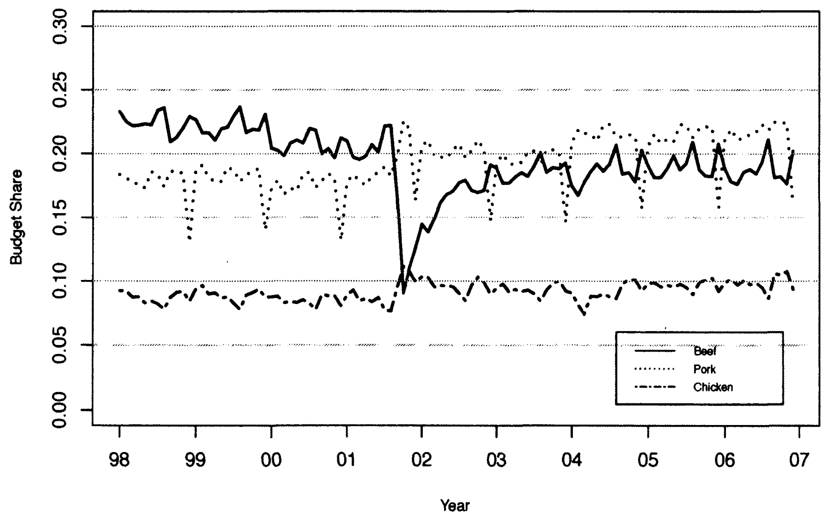

Figure 1: Plot of budget share data

4.1 Estimation Results

We estimate parameters of MS-AIDS model (2.1) with the intercept parameters in (2.4). The Gibbs sampling algorithm is run so that the first 5,000 samples are discarded as burn-in and then the next 25,000 samples are recorded. The prior distributions are parameterized by setting $\boldsymbol{\mu} = \mathbf{0}$, $\boldsymbol{V} = 10^4 \boldsymbol{I}_{45}$, $\nu_j = 10$, $\boldsymbol{\Lambda}_j = 10^{-3}\boldsymbol{I}_3$ $(j = 1, 2)$, $u_{11} = u_{22} = 5$, and $u_{12} = u_{21} = 2$. We also restrict a priori that, in our observed data, at least 40% of observations lie in each regime in order to avoid identification problem within the Gibbs sampling algorithm.

In this study, we examine the following four models: model 1 only includes intercept parameter $\bar{\alpha}_{i,s_t}$. Model 2 includes seasonal effects on August and December into model 1, and model 3 adds a habit effect into model 2. Finally, model 4 further incorporates a trend effect into model 3.

Model 1 $\alpha_{i,s_t} = \bar{\alpha}_{i,s_t}$

Model 2 $\alpha_{i,s_t} = \bar{\alpha}_{i,s_t} + \delta_{1,i} d_{1,t} + \delta_{2,i} d_{2,t}$

Model 3 $\alpha_{i,s_t} = \bar{\alpha}_{i,s_t} + \delta_{1,i} d_{1,t} + \delta_{2,i} d_{2,t} + \sum_{j=1}^{N} \phi_{ij} \bar{w}_{j,t-1}$

Model 4 $\alpha_{i,s_t} = \bar{\alpha}_{i,s_t} + \nu_{i,s_t} t + \delta_{1,i} d_{1,t} + \delta_{2,i} d_{2,t} + \sum_{j=1}^{N} \phi_{ij} \bar{w}_{j,t-1}$

Table 1 shows the logarithmic marginal likelihood of model i, log-ML$_i$ $(i = 1, 2, 3, 4)$ as diagonal elements and logarithmic Bayes factors, log-BF$_{ij}$ for model i against model j as off-diagonal elements. To obtain the marginal likelihoods for candidate models, we use the method proposed by Newton and Raftery (1994). Although models 3 and 4 have large logarithmic marginal likelihoods relative to the other models, logarithmic Bayes factor for model 4 against model 3, log-BF$_{43}$ $(= 1.295)$ indicates "positive" (Kass and Raftery,

Table 1: Log-Marginal Likelihood and Log-Bayes Factor

	Model1	Model2	Model3	Model4
Model1	1098.416	—	—	—
Model2	18.927	1117.343	—	—
Model3	62.157	43.230	1160.573	—
Model4	63.452	44.525	1.295	1161.868

Table 2: Estimated Parameters of MS-AIDS model in Regime1

	Mean	SD	2.5%	50%	97.5%	CD
$\bar{\alpha}_1$	0.2866	0.1262	0.0416	0.2850	0.5373	0.6055
$\bar{\alpha}_2$	0.6046	0.0925	0.4118	0.6074	0.7780	-0.3348
$\bar{\alpha}_3$	0.3285	0.0794	0.1694	0.3290	0.4824	-0.7486
γ_{11}	0.0245	0.0613	-0.1131	0.0350	0.1229	-0.0015
γ_{12}	-0.0039	0.0437	-0.0701	-0.0129	0.0951	0.3383
γ_{13}	-0.0549	0.0335	-0.1315	-0.0534	0.0060	-0.4092
γ_{22}	0.0335	0.0473	-0.0693	0.0368	0.1174	-0.1208
γ_{23}	-0.0110	0.0403	-0.0947	-0.0096	0.0653	-0.7934
γ_{33}	0.0891	0.0480	-0.0020	0.0879	0.1892	0.7943
β_1	-0.0301	0.0265	-0.0817	-0.0301	0.0226	-0.6585
β_2	-0.0936	0.0193	-0.1320	-0.0936	-0.0553	-0.0383
β_3	-0.0195	0.0170	-0.0524	-0.0198	0.0147	0.7026
σ_{11}^2	0.000061	0.000013	0.000041	0.000060	0.000091	1.3920
σ_{12}	-0.000003	0.000008	-0.000018	-0.000003	0.000013	-0.5299
σ_{13}	-0.000008	0.000007	-0.000022	-0.000008	0.000006	-0.1079
σ_{22}^2	0.000035	0.000008	0.000024	0.000034	0.000053	0.4690
σ_{23}	0.000006	0.000005	-0.000004	0.000006	0.000016	0.2985
σ_{33}^2	0.000029	0.000006	0.000019	0.000028	0.000044	1.1468

1995, p.777) evidence in favor of model 4. Therefore we conclude that model 4 fits the data best.

Tables 2 and 3 show the results of parameters for beef, pork and chicken. The parameters for fish are estimated from the adding-up condition (2.3a). Tables 2 and 3 show the the posterior means, posterior standard deviations (SD), 95% credible intervals, and Geweke's convergence diagnostic statistics (CD) for all parameters in MS-AIDS model (2.1) and variance-covariance matrices in regimes 1 and 2. To carry out the Geweke's convergence diagnostic, we used the first 10% and last 50% of the recorded simulated data and Tables 2 and 3 show that all parameters pass the Geweke's convergence diagnostic at 5% significant level. In the Bayesian framework, if a 95% credible interval does not include zero, estimated parameters are interpreted as the significant parameters. Thus $\bar{\alpha}_1$, $\bar{\alpha}_2$, $\bar{\alpha}_3$, β_2, σ_{11}^2, σ_{22}^2, and σ_{33}^2 in Table 2 and $\bar{\alpha}_2$, $\bar{\alpha}_3$, β_2, σ_{11}^2, σ_{12}, σ_{13}, σ_{22}^2, and σ_{33}^2 in Table 3 are regarded significantly different from zero. These parameters in MS-AIDS model (2.1) are used to calculate the price and expenditure elasticities.

To compare our proposed Bayesian estimation with the ML estimation proposed in

Table 3: Estimated Parameters of MS-AIDS model in Regime2

	Mean	SD	2.5%	50%	97.5%	CD
$\bar{\alpha}_1$	0.1926	0.1858	-0.1418	0.1817	0.5981	-0.4185
$\bar{\alpha}_2$	0.6542	0.1255	0.3993	0.6593	0.8891	-0.4349
$\bar{\alpha}_3$	0.2838	0.1017	0.0812	0.2843	0.4830	-0.6659
γ_{11}	0.1032	0.0682	-0.0267	0.1056	0.2270	-0.1384
γ_{12}	-0.0484	0.0355	-0.1164	-0.0489	0.0185	0.4018
γ_{13}	-0.0560	0.0299	-0.1154	-0.0555	0.0011	0.9316
γ_{22}	0.0432	0.0417	-0.0368	0.0425	0.1272	0.0669
γ_{23}	0.0236	0.0343	-0.0436	0.0234	0.0923	-0.3294
γ_{33}	0.0628	0.0399	-0.0146	0.0620	0.1429	-1.1898
β_1	-0.0329	0.0385	-0.1131	-0.0318	0.0392	0.8659
β_2	-0.0949	0.0246	-0.1421	-0.0954	-0.0454	0.5194
β_3	-0.0112	0.0210	-0.0529	-0.0111	0.0301	0.0316
σ_{11}^2	0.000161	0.000045	0.000096	0.000154	0.000267	-0.7579
σ_{12}	-0.000033	0.000018	-0.000074	-0.000031	-0.000004	0.2701
σ_{13}	-0.000029	0.000015	-0.000062	-0.000027	-0.000003	0.3091
σ_{22}^2	0.000045	0.000010	0.000030	0.000044	0.000069	-1.0053
σ_{23}	0.000004	0.000008	-0.000011	0.000004	0.000020	-0.0595
σ_{33}^2	0.000042	0.000010	0.000027	0.000041	0.000065	-1.3308

Allais and Nichèle (2007) and employed in Kabe and Kanazawa (2012), we calculate the mean squared errors (MSEs) for estimated budget shares. The MSEs of Bayesian estimation are evaluated by the posterior means of estimates of budget shares generated within the Gibbs sampler. The results of MSEs with respect to our proposed Bayesian estimation (Bayes) and ML estimation (MLE) are given in Table 4. Our Bayesian estimation improves the MSEs for all products over ML estimation. This result reflects the goodness of fit to the budget share data of our Bayesian estimation.

Table 4: Mean squared errors (MSEs)

	Beef	Pork	Chicken	Fish
Bayes	0.805×10^{-4}	0.172×10^{-4}	0.152×10^{-4}	0.603×10^{-4}
MLE	1.177×10^{-4}	0.264×10^{-4}	0.158×10^{-4}	0.657×10^{-4}

Figure 2 plots the probability of being regime 2 and budget share data of beef and pork from January 1998 through December 2006 under the proposed Bayesian method. We calculate the probability $\Pr\{s_t = 2\}$. In Figure 2, regime shift from $s_t = 1$ to $s_t = 2$ is observed at the timing of first BSE case in Japan in September 2001 and then the probability $\Pr\{s_t = 2\}$ gradually declines until the end of 2003 along with increase in budget share of beef. With the timing of first BSE case in U.S. in December 2003, we observe a high probability of being regime 2 once again. Since then, structure of budget share tends to stay in regime 2.

The first regime shift in Figure 2 reflects the switching of consumers' preference from beef to pork triggered by the BSE scare in Japan in September 2001. The second regime shift after the first BSE discovery in U.S. in December 2003 might have arisen due to the ban on import of American beef. Since the ban on importing U.S. beef led to the shortage of beef supply in the domestic meat market, consumers may have been forced to purchase more pork instead of beef.

Figure 3 shows the results of probability of being regime 2, $\Pr(s_t = 2|\boldsymbol{\Omega}_t, \widehat{\boldsymbol{\Theta}})$ under the ML estimation. We estimated the probability from the Hamilton filter using data set obtained through time t, $\boldsymbol{\Omega}_t$, and ML estimates $\widehat{\boldsymbol{\Theta}}$ in model 4 (see Kabe and Kanazawa, 2012). The regime shift at the timing of first BSE case in Japan in September 2001 is observed in Figure 3. Nevertheless, when compared with the result of regime shift in Figure 2, the probability's gradual decline due to the recovery of beef budget share following the first BSE case in Japan observed in Figure 2 no longer can be observed in Figure 3. Unlike Figure 2, we cannot identify the regime shift at the timing of first BSE case in U.S. in December 2003 in Figure 3.

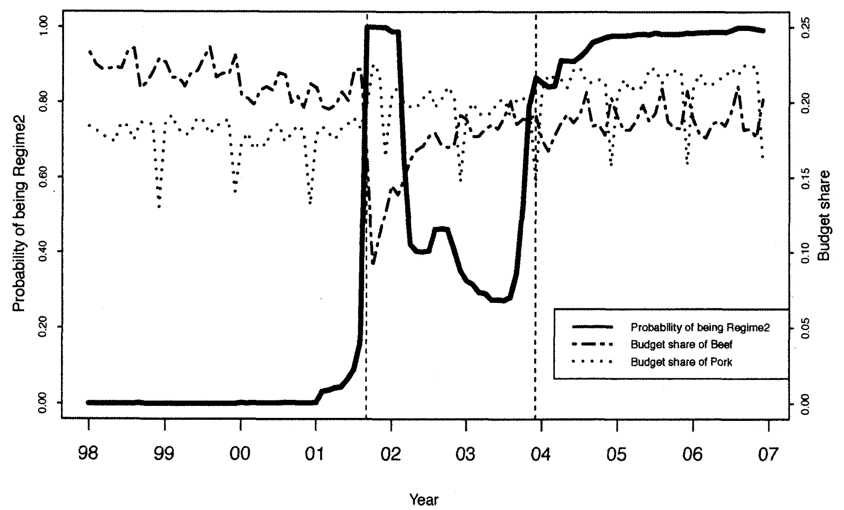

Figure 2: Probability of being regime 2, $\Pr\{s_t = 2\}$ and budget share data of beef and pork under the proposed Bayesian estimation. Two vertical dashed lines indicate the first BSE case in Japan on September 2001 and the first BSE case in U.S. on December 2003.

We calculate the average budget share of i-th product at regime $s_t = j$ as

$$\bar{w}_{i,s_t=j} = \frac{\sum_{t=1}^{T} \mathbf{1}\{s_t = j\}\bar{w}_{it}}{\sum_{t=1}^{T} \mathbf{1}\{s_t = j\}}.$$

Table 5 shows that regime 1 is characterized by a higher beef budget share relative to that of pork, while regime 2 is characterized by the reversal of these two budget shares.

Since substitution occurs mostly between beef and pork in regimes 1 and 2 (see Table 5), we focus on the price and expenditure elasticities for beef and pork. We calculate

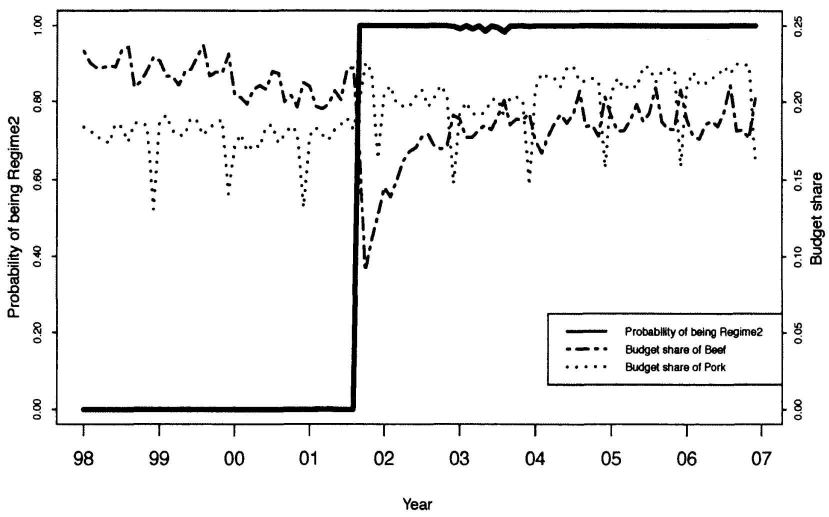

Figure 3: Probability of being regime 2, $\Pr(s_t = 2|\boldsymbol{\Omega}_t, \widehat{\boldsymbol{\Theta}})$ and budget share data of beef and pork under ML estimation.

Table 5: Posterior mean of average budget share

	Regime 1	Regime 2
Beef	0.2075	0.1799
Pork	0.1825	0.2071
Chicken	0.0878	0.0961
Fish	0.5222	0.5170

the Marshallian price elasticity η^P_{ij,s_t} and expenditure elasticity η^E_{i,s_t} at regime s_t for each 25,000 samples generated via Gibbs sampler as

$$\eta^P_{ij,s_t} = -\kappa_{ij} + \frac{\gamma_{ij,s_t}}{\bar{w}_{i,s_t}} - \frac{\beta_{i,s_t}}{\bar{w}_{i,s_t}}\left[\alpha_{j,s_t} + \sum_{k=1}^{N}\gamma_{kj,s_t}\log\bar{p}_{k,s_t}\right], \tag{4.1}$$

$$\eta^E_{i,s_t} = \frac{\beta_{i,s_t}}{\bar{w}_{i,s_t}} + 1, \tag{4.2}$$

where $\kappa_{ij} = 1$ for $i = j$ and $\kappa_{ij} = 0$ for $i \neq j$, and $\bar{p}_{k,s_t}$ is an average price at regime s_t.

In Table 6, we show the posterior means and 95% credible intervals of price and expenditure elasticities for beef and pork. Although own-price elasticities of pork have significant negative impacts in both regimes, own-price elasticity of beef in regime 2 includes zero within the 95% credible interval. Since American beef was banned and in short supply then, beef prices tended to increase in regime 2. Hence this price inelastic beef purchasing behavior in regime 2 in Table 6 implies that those who had kept purchasing beef in regime 2 did so regardless of its price.

Table 6: Price elasticities and Expenditure elasticities

Regime1	Price (η_{ij}^P)		Expenditure (η_{ij}^E)
	Beef	Pork	
Beef	-0.8171	0.0600	0.8552
	(-1.4682, -0.3440)	(-0.2839, 0.5459)	(0.6088, 1.1092)
Pork	0.1483	-0.5142	0.4874
	(-0.2299, 0.7208)	(-1.0636, -0.0365)	(0.2783, 0.6939)

Regime2	Price (η_{ij}^P)		Expenditure (η_{ij}^E)
	Beef	Pork	
Beef	-0.3357	-0.1679	0.8168
	(-1.0458, 0.3696)	(-0.5573, 0.2385)	(0.3694, 1.2182)
Pork	-0.0961	-0.5050	0.5421
	(-0.4100, 0.2209)	(-0.9362, -0.0620)	(0.3172, 0.7797)

1) 95% credible interval in parentheses

5 Conclusion

In this paper, we proposed the Bayesian estimation for MS-AIDS model proposed by Allais and Nichèle (2007) and illustrated the applicability of our proposed method via real data. The proposed Bayesian estimation has some important advantages. First, it enables us to avoid the singularity problem suggested in Hamilton (1990, 1991). In the Bayesian framework, we can use conjugate prior distributions to incorporate the prior information about variance-covariance matrices in advance. On the other hand, ML estimation via numerical optimization methods (e.g., Newton-Raphson method) has to depend on sensible selection of initial values of parameters to avoid singularity points on the parameter space. Second, our proposed Bayesian estimation by design ensures that transition probabilities be located between zero and one by generating the samples from the beta distributions within the Gibbs sampler. Third, there is no need to calculate the score functions of log-likelihood, unlike ML estimation which employed in Allais and Nichèle (2007) and Kabe and Kanazawa (2012), is computationally very intensive. In our Bayesian estimation, posterior distributions of parameters are expressed as the standard formula (e.g., multivariate normal and inverse Wishart distributions). Thus each of parameters can be easily simulated via Gibbs sampler.

In the empirical study on the Japanese meat market, we found that our Bayesian estimation improves the mean squared errors for all meat products compared with the ML estimation. Moreover we found the regime shift in the budget shares of meat products in Figures 2 depicts much more sophisticated and realistic picture of regime transition than Figure 3. Specifically, in Figure 2, probability of being regime 2, $\Pr\{s_t = 2\}$, estimated via Bayesian estimation shows the regime shifts at the timing of first reported cases of BSE both in Japan in September 2001 and also in U.S. in December 2003. On the other hand, Figure 3 shows a single regime shift at the timing of first BSE case in Japan. Since

ML parameter estimates are given as the point solutions, the probability of being regime 2, $\Pr(s_t = 2|\Omega_t, \widehat{\Theta})$ in Figure 3 ignores the uncertainty about parameters and making the probability $\Pr(s_t = 2|\Omega_t, \widehat{\Theta})$ in Figure 3 closer to zero or one. Perhaps as Scott (2002, p.345) observed, ignoring uncertainly about the parameter may have contributed to such result.

Finally, we discuss the further extension of MS-AIDS model. Several studies have extended the Hamilton (1989)'s Markov-switching model. In particular, they focused on a useful modification of transition probabilities. For example, Diebold et al. (1994) introduced the time-varying transition probabilities into the Markov-switching model, and also they allowed the transition probabilities to evolve as logistic functions of economic variables. Alternatively, Markov-switching model assumes that latent variables controlling regime shifts are exogenous. Kim et al. (2008) relaxed this exogenous regime-switching assumption and proposed a Markov-switching regression model with endogenous regime switching. The extensions to MS-AIDS model in these directions will be interesting for future research.

References

Allais, O., and V. Nichèle (2007) 'Capturing structural changes in french meat and fish demand over the period 1991–2002.' *European review of agricultural economics* 34(4), 517–538

Banks, J., R. Blundell, and A. Lewbel (1997) 'Quadratic engel curves and consumer demand.' *Review of Economics and Statistics* 79(4), 527–539

Carter, C., and R. Kohn (1994) 'On gibbs sampling for state space models.' *Biometrika* 81(3), 541–553

Chib, S. (1996) 'Calculating posterior distributions and modal estimates in markov mixture models.' *Journal of Econometrics* 75(1), 79–97

Christensen, L. R., D. W. Jorgenson, and L. J. Lau (1975) 'Transcendental logarithmic utility functions.' *American Economic Review* 65(3), 367–383

Cooper, R. J., and K. R. McLaren (1992) 'An empirically oriented demand system with improved regularity properties.' *Canadian Journal of Economics* 25(3), 652–668

Deaton, A., and J. Muellbauer (1980) 'An almost ideal demand system.' *The American Economic Review* 70(3), 312–326

Diebold, F. X., J. H. Lee, and G. Weinbach (1994) 'Regime switching with time-varying transition probabilities.' in C. Hargreaves (ed.), Nonstationary Time Series Analysis and Cointegration.(Advanced Texts in Econometrics, CWJ Granger and G. Mizon, eds.), 283-302. Oxford: Oxford University Press.

Hamilton, J. D. (1989) 'A new approach to the economic analysis of nonstationary time series and the business cycle.' *Econometrica: Journal of the Econometric Society* 57(2), 357–384

_ (1990) 'Analysis of time series subject to changes in regime.' *Journal of Econometrics* 45(1-2), 39–70

_ (1991) 'A quasi-bayesian approach to estimating parameters for mixtures of normal distributions.' *Journal of Business & Economic Statistics* pp. 27–39

Ishida, T., N. Ishikawa, and M. Fukushige (2006) 'Impact of bse and bird flu on consumers' meat demand in japan.' *DISCUSSION PAPER SERIES in Osaka university*

Ishida, T., N. Ishikawa, and M. Fukushige (2010) 'Impact of bse and bird flu on consumers' meat demand in japan.' *Applied Economics* 42(1), 49–56

Jin, H. J., and W. W. Koo (2003) 'The effect of the bse outbreak in japan on consumers' preferences.' *European Review of Agricultural Economics* 30(2), 173–192

Kabe, S., and Y. Kanazawa (2012) 'Another view of impact of bse crisis in japanese meat market through the almost ideal demand system model with markov switching.' *Department of Social Systems and Management Discussion Paper Series No.1288 (University of Tsukuba), forthcoming in Applied Economics Letters (Accepted on December 2011).* URL: http://www.sk.tsukuba.ac.jp/SSM/libraries/list1276.php

Kass, R. E., and A. E. Raftery (1995) 'Bayes factors.' *The Journal of the American Statistical Association* 90(430), 773–795

Kim, C. J., J. Piger, and R. Startz (2008) 'Estimation of markov regime-switching regression models with endogenous switching.' *Journal of Econometrics* 143(2), 263–273

Moosa, I. A., and J. L. Baxter (2002) 'Modeling the trend and seasonal within an aids model of demand for alcoholic beverage in the united kingdom.' *Journal of Applied Econometrics* 17(2), 95–106

Newton, M. A., and A. E. Raftery (1994) 'Approximate bayesian inference with the weighted likelihood bootstrap.' *Journal of the Royal Statistical Society, Series B* 56(1), 3–48

Ohtani, K., and S. Katayama (1986) 'A gradual switching regression model with autocorrelated errors.' *Economics Letters* 21(2), 169–172

Peterson, H. H., and Y. J. K. Chen (2005) 'The impact of bse on japanese retail meat demand.' *Agribusiness* 21(3), 313–327

Rickertsen, K. (1996) 'Structural change and the demand for meat and fish in norway.' *European Review of Agricultural Economics* 23(3), 316–330

Scott, S. L. (2002) 'Bayesian methods for hidden markov models.' *Journal of the American Statistical Association* 97(457), 337–351

Theil, H. (1965) 'The information approach to demand analysis.' *Econometrica* 33(1), 67–87